

LIPSCHITZ CONTINUITY OF THE VALUE FUNCTION IN MIXED-INTEGER OPTIMAL CONTROL PROBLEMS*

MARTIN GUGAT[†], FALK M. HANTE[†]

ABSTRACT. We study the optimal value function for control problems on Banach spaces that involve both continuous and discrete control decisions. For problems involving semilinear dynamics subject to mixed control inequality constraints, one can show that the optimal value depends locally Lipschitz continuously on perturbations of the initial data and the costs under rather natural assumptions. We prove a similar result for perturbations of the initial data, the constraints and the costs for problems involving linear dynamics, convex costs and convex constraints under a Slater-type constraint qualification. We show by an example that these results are in a sense sharp.

1. INTRODUCTION

In this paper we address the robustness of solutions to optimal control problems that involve both continuous-valued and discrete-valued control decisions to steer solutions of a differential equation such that an associated cost is minimized. This problem class includes in particular optimal control of switched systems [20, 21], but also optimization of systems with coordinated activation of multiple actuators, for example, at different locations in space for certain distributed parameter systems [12, 11]. In analogy to mixed-integer programming we call such problems mixed-integer optimal control problems. Algorithms to compute solutions to such problems are discussed in [8, 16, 11, 17, 18, 15]. From a theoretical point of view, but also for a reliable application of such algorithms, the robustness of the solution with respect to perturbation of data in the problem is essential, for instance, in the case of uncertain initial data. We consider the robustness of the optimal value because this is the criterion determining the control decision. Moreover, we understand robustness in the sense that we consider the regularity of the optimal value as a function of the problem parameters.

For continuous optimization problems many sensitivity results are available, see [3, 13]. In particular certain regularity assumptions and constraint qualifications guarantee the continuity of the optimal value function, see [7, 9]. In the context of mixed-integer programming, in general, the main difficulty is that the admissible set consists of several connected components and jumps in the optimal value as function of the problem parameters can occur if due to parameter changes connected components of the feasible set vanish. In mixed-integer linear programming with bounded feasible sets, the continuity of the value function is therefore equivalent to existence of a Slater-point [19]. For mixed-integer convex programs, constraint qualifications are given in [10] which yield the existence of one-sided directional

Date: December 29, 2016.

Key words and phrases. Parametric optimal control; parametric switching control; parametric optimization; sensitivity; mixed-integer optimal control problems; optimal value function; Lipschitz continuity.

* The article is published in *Math. Control Signals Syst.* (2017) 29:3.

[†] Lehrstuhl für Angewandte Mathematik 2, Department Mathematik, Friedrich-Alexander Universität Erlangen-Nürnberg, {falk.hante,martin.gugat}@fau.de.

derivatives of the value function and hence its Lipschitz continuity. For optimal control problems in general, it is well known that one cannot expect more regularity of the optimal value function than Lipschitz continuity. The following example is an adaption of a classical one saying that this is also true for integer, and hence mixed-integer controlled systems.

Example 1. For some $t_f > 0$ and $\lambda \in \mathbb{R}$, consider the problem

$$\left. \begin{array}{l} \text{minimize } y(t_f) \text{ subject to} \\ \dot{y}(t) = v(t)y(t), \text{ for a.e. } t \in (0, t_f), \quad y(0) = \lambda \\ y(t) \in \mathbb{R}, \quad v(t) \in \{0, 1\} \text{ for a.e. } t \in (0, t_f). \end{array} \right\} \quad (1)$$

The optimal value function $\nu(\lambda) = \inf\{y(t_f; \lambda) : v \in L^\infty(0, t_f; \{0, 1\})\}$ can easily be seen to be

$$\nu(\lambda) = \begin{cases} e^{t_f} \lambda, & \lambda < 0, \\ \lambda & \lambda \geq 0, \end{cases}$$

which is Lipschitz continuous, but not differentiable in $\lambda = 0$.

For semilinear mixed-integer optimal control problems, we show below that for parametric initial data as in the example, local Lipschitz continuity of the optimal value function can indeed be guaranteed for a rather general setting without imposing a Slater-type condition. Similar results are well-known in the classical Banach or Hilbert space case without mixed control constraints [5, 1]. Further, we analyze parametric control constraints and parametric cost functions for convex programs. For this case, we formulate a Slater-type condition guaranteeing again the local Lipschitz continuity of the optimal value function. Finally, for convex programs, we can combine both results to obtain local Lipschitz continuity jointly for parametric initial data, control constraints and cost functions.

2. SETTING AND PRELIMINARIES

Let Y be a Banach space, U be a complete metric space, \mathcal{V} be a finite set, and $f: [t_0, t_f] \times Y \times U \times \mathcal{V} \rightarrow Y$. We consider the control system

$$\dot{y}(t) = Ay(t) + f(t, y(t), u(t), v(t)), \quad t \in (t_0, t_f) \text{ a.e.}, \quad (2)$$

where $[t_0, t_f]$ is a finite time horizon with $t_0 < t_f$, $A: D(A) \rightarrow Y$ is a generator of a strongly continuous semigroup $\{T(t)\}_{t \geq 0}$ of bounded linear operators on Y , and where $u: [t_0, t_f] \rightarrow U$ and $v: [t_0, t_f] \rightarrow \mathcal{V}$ are two independent measurable control functions. Throughout the paper we consider the Lebesgue-measure. Our main concern will be the confinement that the control v only takes values from a finite set. Without loss of generality, we may identify \mathcal{V} with a set of integers $\{0, 1, \dots, N-1\}$ and, in analogy to mixed-integer programming, we refer to (2) as a *mixed-integer control system*, where u represents ordinary controls and v integer controls. Let $U_{[t_0, t_f]}$ be a Banach subspace of measurable ordinary control functions $u: [t_0, t_f] \rightarrow U$ and let $V_{[t_0, t_f]}$ be the set of measurable integer control functions $v: [t_0, t_f] \rightarrow \mathcal{V}$. By the assumed finiteness of \mathcal{V} we actually have $V_{[t_0, t_f]} = L^\infty(t_0, t_f; \mathcal{V})$.

Let Λ be a Banach space and consider (2) subject to a parametric initial condition

$$y(t_0) = y_0(\lambda), \quad (3)$$

where $y_0(\lambda)$ is an initial state in Y parametrized by $\lambda \in \Lambda$.

The separation of the control in u and v and the inherent integer confinement of the latter control lets us formulate parametric control constraints of the mixed form

$$g_k^v(\lambda, u, t) \leq 0, \quad k = 1, \dots, M, \quad t \in [t_0, t_f] \quad (4)$$

where $M \in \mathbb{N}$ and, for every $v \in V_{[t_0, t_f]}$, the functions $g_1^v, \dots, g_M^v: \Lambda \times U_{[t_0, t_f]} \times [t_0, t_f] \rightarrow \mathbb{R}$ are given. These constraints can for example model anticipating control

restrictions, where a decision represented by v at an earlier time limits control decisions for u at different times. We discuss an example in Section 5. In cases without mixed control constraints, we set $M = 0$.

Definition 1. For fixed $\lambda \in \Lambda$, let $W_{[t_0, t_f]}(\lambda)$ denote the set of all admissible controls

$$W_{[t_0, t_f]}(\lambda) := \{(u, v) \in U_{[t_0, t_f]} \times V_{[t_0, t_f]} : g_k^v(\lambda, u, t) \leq 0, k = 1, \dots, M, t \in [t_0, t_f]\}. \quad (5)$$

Moreover, we say that $y: [t_0, t_f] \rightarrow Y$ is a *solution of the mixed-integer control system* if there exists an admissible pair of controls $(u, v) \in W_{[t_0, t_f]}(\lambda)$ such that $y \in C([t_0, t_f]; Y)$ satisfies the integral equation

$$y(t) = T(t - t_0)y(t_0) + \int_{t_0}^t T(t - s)f(s, y(s), u(s), v(s)) ds, t \in [t_0, t_f] \quad (6)$$

and (3) holds. Let $\mathcal{S}_{[t_0, t_f]}(\lambda)$ denote the set of all such solutions y defined on $[t_0, t_f]$. For any $y \in \mathcal{S}_{[t_0, t_f]}(\lambda)$, we denote by $y = y(\cdot; y_0(\lambda), u, v)$ the dependency of y on $y_0(\lambda)$, u and v if needed.

According to Definition 1, $\mathcal{S}_{[t_0, t_f]}(\lambda)$ consists of the mild solutions of equation (2) and covers in an abstract sense many evolution problems involving linear partial differential operators [14]. In particular, the mild solutions coincide with the usual concept of weak solutions in case of linear parabolic partial differential equations on reflexive Y with distributed control where A arises from a time-invariant variational problem [2]. For an example, see Section 5.

In conjunction with the mixed-integer control system we consider a cost function $\varphi: \Lambda \times C([t_0, t_f]; Y) \times U_{[t_0, t_f]} \times V_{[t_0, t_f]} \rightarrow \mathbb{R} \cup \{\infty\}$ and define the *mixed-integer optimal control problem* with parameter λ as

$$\left. \begin{aligned} &\text{minimize } \varphi(\lambda, y, u, v) \text{ subject to} \\ &\dot{y}(t) = Ay(t) + f(t, y(t), u(t), v(t)), t \in (t_0, t_f) \text{ a.e.}, \\ &y(0) = y_0(\lambda), \\ &g_k^v(\lambda, u, t) \leq 0 \text{ for all } t \in [t_0, t_f], k = 1, \dots, M, \\ &y \in C([t_0, t_f]; Y), u \in U_{[t_0, t_f]}, v \in V_{[t_0, t_f]}. \end{aligned} \right\} \quad (7)$$

We will study the corresponding *optimal value* $\nu(\lambda) \in \mathbb{R} \cup \{\pm\infty\}$ given by

$$\left. \begin{aligned} &\nu(\lambda) = \inf \{ \varphi(\lambda, y, u, v) : \\ &\dot{y}(t) = Ay(t) + f(t, y(t), u(t), v(t)), t \in (t_0, t_f) \text{ a.e.}, \\ &y(0) = y_0(\lambda), \\ &g_k^v(\lambda, u, t) \leq 0 \text{ for all } t \in [t_0, t_f], k = 1, \dots, M, \\ &y \in C([t_0, t_f]; Y), u \in U_{[t_0, t_f]}, v \in V_{[t_0, t_f]} \} \end{aligned} \right\} \quad (8)$$

in its dependency on the parameter λ .

For the mixed-integer control system, we will impose the following assumptions.

Assumption 1. The map $f: [t_0, t_f] \times Y \times U \times \{v\} \rightarrow Y$ is continuous for all $v \in \mathcal{V}$. Moreover, there exists a function $k \in L^1(t_0, t_f)$ such that for all $(u, v) \in W_{[t_0, t_f]}$, $y_1, y_2 \in Y$ and for almost every $t \in (t_0, t_f)$

- (i) $|f(t, y_1, u(t), v(t)) - f(t, y_2, u(t), v(t))| \leq k(t)|y_1 - y_2|$
- (ii) $|f(t, 0, u(t), v(t))| \leq k(t)$.

In particular, under these assumptions, the integral in (6) is well-defined in the Lebesgue-Bochner sense and from the theory of abstract Cauchy problems [14] we obtain a solution y in $C([0, t_f]; Y)$ for all $y_0 \in Y$, $u \in U_{[t_0, t_f]}$ and $v \in V_{[t_0, t_f]}$. Moreover, the strong continuity of $T(\cdot)$ and the Gronwall inequality yield the following solution properties.

Lemma 1. *Under the Assumptions 1, there exist constants $\gamma \geq 0$ and $w_0 \geq 0$ such that for all $\lambda_1, \lambda_2 \in \Lambda$, setting $y_i = y(\cdot; y_0(\lambda_i), u, v) \in \mathcal{S}_{[t_0, t_f]}(\lambda_i)$ for $i \in \{1, 2\}$, for all $t \in [t_0, t_f]$ it holds $\|T(t)\| \leq \gamma \exp(w_0(t - t_0))$,*

$$|y_i(t)| \leq C(t)(1 + |y_0(\lambda_i)|), \quad i \in \{1, 2\}, \quad (9)$$

and

$$|y_1(t) - y_2(t)| \leq C(t)|y_0(\lambda_1) - y_0(\lambda_2)| \quad (10)$$

with $C(t) = \gamma \exp\left(w_0(t - t_0) + \gamma \int_{t_0}^t k(s) ds\right)$.

For the cost function and control constraints, we will impose the following assumptions.

Assumption 2. The function $\varphi: \Lambda \times C([t_0, t_f]; Y) \times U_{[t_0, t_f]} \times V_{[t_0, t_f]} \rightarrow \mathbb{R}$ is continuous and, for every $v \in V_{[t_0, t_f]}$, the functions $g_1^v, \dots, g_M^v: \Lambda \times U_{[t_0, t_f]} \times [t_0, t_f] \rightarrow \mathbb{R}$ are such that the set of admissible controls $W_{[t_0, t_f]}(\lambda)$ is not empty for all $\lambda \in \Lambda$.

In particular, under Assumptions 1 and 2, for every $\lambda \in \Lambda$, the set $\mathcal{S}_{[t_0, t_f]}(\lambda)$ is non-empty. Moreover, one obtains local Lipschitz continuity of the value function if the perturbation parameter λ acts Lipschitz continuously on φ and y_0 by similar arguments as in a classical Banach or Hilbert space case [5, 1].

Theorem 1. *Under the Assumptions 1 and 2, suppose that the constraint functions g_1^v, \dots, g_M^v are independent of λ . Let $\bar{\lambda}$ be some fixed parameter in Λ and assume that for some bounded neighborhood $B(\bar{\lambda})$ of $\bar{\lambda}$ and some constant L_0*

$$|y_0(\lambda_1) - y_0(\lambda_2)| \leq L_0 |\lambda_1 - \lambda_2|, \quad \lambda_1, \lambda_2 \in B(\bar{\lambda}). \quad (11)$$

Moreover, let $K = \sup_{\lambda \in B(\bar{\lambda})} |y_0(\lambda)|$ and assume that for some constant L_φ

$$|\varphi(\lambda_1, y, u, v) - \varphi(\lambda_2, \bar{y}, u, v)| \leq L_\varphi(|y - \bar{y}| + |\lambda_1 - \lambda_2|) \quad (12)$$

for all $(u, v) \in W_{[t_0, t_f]}$, y, \bar{y} such that $\max\{|y|, |\bar{y}|\} \leq C(t_f)(1 + K)$ and $\lambda_1, \lambda_2 \in B(\bar{\lambda})$, where $C(t)$ is the bound from Lemma 1. Then there exists a constant \hat{L}_ν such that

$$|\nu(\lambda_1) - \nu(\lambda_2)| \leq \hat{L}_\nu |\lambda_1 - \lambda_2|, \quad \lambda_1, \lambda_2 \in B(\bar{\lambda}). \quad (13)$$

Proof. Let $\varepsilon > 0$ and $\lambda_1, \lambda_2 \in B(\bar{\lambda})$ be given. Choose $(u_\varepsilon, v_\varepsilon) \in W_{[t_0, t_f]}$ such that

$$\varphi(\lambda_2, \bar{y}_\varepsilon, u_\varepsilon, v_\varepsilon) \leq \nu(\lambda_2) + \varepsilon,$$

where \bar{y}_ε denotes the reference solution $y(\cdot; y_0(\lambda_2), u_\varepsilon, v_\varepsilon) \in \mathcal{S}_{[t_0, t_f]}$. Let y_ε denote the perturbed solution $y(\cdot; y_0(\lambda_1), u_\varepsilon, v_\varepsilon) \in \mathcal{S}_{[t_0, t_f]}$. Lemma 1 and the assumptions yield

$$|y_\varepsilon(t)| \leq C(t)(1 + K), \quad t \in [t_0, t_f],$$

and

$$|y_\varepsilon(t) - \bar{y}_\varepsilon(t)| \leq C(t)L_0|\lambda_1 - \lambda_2|, \quad t \in [t_0, t_f].$$

Hence,

$$\begin{aligned} \varphi(\lambda_1, y_\varepsilon, u_\varepsilon, v_\varepsilon) &\leq \varphi(\lambda_2, \bar{y}_\varepsilon, u_\varepsilon, v_\varepsilon) + |\varphi(\lambda_1, y_\varepsilon, u_\varepsilon, v_\varepsilon) - \varphi(\lambda_2, \bar{y}_\varepsilon, u_\varepsilon, v_\varepsilon)| \\ &\leq \varphi(\lambda_2, \bar{y}_\varepsilon, u_\varepsilon, v_\varepsilon) + L_\varphi(C(t_f)L_0 + 1)|\lambda_1 - \lambda_2|. \end{aligned}$$

Thus

$$\begin{aligned} \nu(\lambda_1) &\leq \varphi(\lambda_1, y_\varepsilon, u_\varepsilon, v_\varepsilon) \leq \varphi(\lambda_2, \bar{y}_\varepsilon, u_\varepsilon, v_\varepsilon) + L_\varphi(C(t_f)L_0 + 1)|\lambda_1 - \lambda_2| \\ &\leq \nu(\lambda_2) + \varepsilon + L_\varphi(C(t_f)L_0 + 1)|\lambda_1 - \lambda_2|. \end{aligned}$$

Letting $\varepsilon \rightarrow 0$ from above gives an upper bound $\nu(\lambda_1) \leq \nu(\lambda_2) + \hat{L}_\nu |\lambda_1 - \lambda_2|$ with

$$\hat{L}_\nu = L_\varphi(C(t_f)L_0 + 1). \quad (14)$$

Interchanging the roles of λ_1 and λ_2 yields the claim. \square

In the subsequent section, we will obtain a similar result concerning the perturbation of the functions g_1^v, \dots, g_M^v and the cost function φ under additional structural hypothesis and a constraint qualification.

3. PERTURBATION OF THE CONSTRAINTS FOR CONVEX PROBLEMS

In this section, we show that under a Slater-type condition the optimal value $\nu(\lambda)$ of the mixed-integer optimal control problem (7) in the case of a convex cost function and linear dynamics is locally Lipschitz continuous as a function of a parameter λ acting on the control constraints g_1^v, \dots, g_M^v and the cost function φ . We need the following

Assumption 3. The map $(y, u) \mapsto f(t, y, u, v)$ is linear and the map $(y, u) \mapsto \varphi(\lambda, y, u, v)$ is convex. Moreover, the function φ is Lipschitz continuous with respect to λ in the sense that

$$|\varphi(\lambda_1, y, u, v) - \varphi(\lambda_2, y, u, v)| \leq L_\varphi(|y|, |u|) |\lambda_1 - \lambda_2| \quad (15)$$

with a continuous function $L_\varphi: [0, \infty)^2 \rightarrow [0, \infty)$. For all $k = 1, \dots, M$, the maps $u \mapsto g_k^v(\lambda, u, t)$ are convex, the maps $(u, t) \mapsto g_k^v(\lambda, u, t)$ are continuous and the functions g_k^v are Lipschitz continuous with respect to λ in the sense that for all $t \in [t_0, t_f]$

$$|g_k^v(\lambda_1, u, t) - g_k^v(\lambda_2, u, t)| \leq L_g(|u|) |\lambda_1 - \lambda_2| \quad (16)$$

with a continuous function $L_g: [0, \infty) \rightarrow [0, \infty)$.

Under the Assumptions 1–3 and assuming that y_0 is independent of λ , we have for each parameter $\lambda \in \Lambda$ the mixed-integer optimal control problem (7) with $y_0(\lambda)$ replaced by a fixed initial state $y_0 \in Y$. Moreover, in this section, $\nu(\lambda)$ denotes the corresponding optimal value function (8) with fixed initial state y_0 . The subsequent analysis is based upon the presentation in [10], where for the finite dimensional case the existence of the one sided derivatives of the optimal value function $\nu(\lambda)$ is shown. For a generalization to the above setting, we first introduce a Slater-type constraint qualification, a dual problem and prove a strong duality result.

Assumption 4. (CQ) For some $\bar{\lambda} \in \Lambda$ and some bounded neighborhood $B(\bar{\lambda}) \subset \Lambda$ of $\bar{\lambda}$ there exists a number $\omega > 0$ such that for all $v \in V_{[t_0, t_f]}$ there is a Slater point $\bar{u}_v \in U$ such that for all $\lambda \in B(\bar{\lambda})$ we have

$$g_k^v(\lambda, \bar{u}_v, t) \leq -\omega \quad \text{for all } t \in [t_0, t_f], \quad k = 1, \dots, M, \quad (17)$$

$$\sup_{v \in V_{[t_0, t_f]}} \sup_{\lambda \in B(\bar{\lambda})} \varphi(\lambda, y(\bar{u}_v, v), \bar{u}_v, v) < \infty \quad (18)$$

and that there exists a number $\underline{\alpha}$ such that for all $\lambda \in B(\bar{\lambda})$ we have

$$\nu(\lambda) \geq \underline{\alpha} \quad (19)$$

and that the set

$$\bigcup_{\lambda_1, \lambda_2 \in B(\bar{\lambda})} \left\{ (u, v) \in U_{[t_0, t_f]} \times V_{[t_0, t_f]} : \begin{aligned} &\varphi(\lambda_1, y(u, v), u, v) \leq \varphi(\lambda_1, y(\bar{u}_v, v), \bar{u}_v, v) + |\lambda_1 - \lambda_2|^2, \\ &g_k^v(\lambda_1, u(t), t) \leq 0, \quad t \in [t_0, t_f], \quad k = 1, \dots, M \end{aligned} \right\} =: \bar{S}(y_0) \quad (20)$$

is bounded.

Note that (19) holds if $\underline{\alpha}$ is a lower bound for the cost function.

Let $C([t_0, t_f])_+^*$ denote the set of positive function of bounded variation on $[t_0, t_f]$. For any controls $v \in V_{[t_0, t_f]}$, $u \in U_{[t_0, t_f]}$ and any $\mu^* \in (C([t_0, t_f])_+^*)^M$ we define the Lagrangian

$$\mathcal{L}_v(\lambda, u, \mu^*) = \varphi(\lambda, y(u, v), u, v) + \sum_{k=1}^M \int_{t_0}^{t_f} g_k^v(\lambda, u, s) d\mu_k^*(s), \quad (21)$$

where the integral is in the Riemann-Stieltjes sense. Further, we define

$$h_v(\lambda, \mu^*) = \inf_{u \in U_{[t_0, t_f]}} \mathcal{L}_v(\lambda, u, \mu^*). \quad (22)$$

Under the constraint qualification (CQ), for all fixed $\lambda \in B(\bar{\lambda})$ and $v \in V_{[t_0, t_f]}$, the classical convex duality theory as presented in [6] implies the strong duality result (see also [9])

$$\sup_{\mu^* \in (C([t_0, t_f])_+^*)^M} h_v(\lambda, \mu^*) = \nu^v(\lambda) \quad (23)$$

where $\nu^v(\lambda)$ denotes the optimal value of the following convex optimal control problem only in the variables y and u

$$\left. \begin{array}{l} \text{minimize } \varphi(\lambda, y, u, v) \text{ subject to} \\ \dot{y}(t) = A y(t) + f(t, y(t), u(t), v(t)), \quad t \in (t_0, t_f) \text{ a.e., } y(0) = y_0, \\ g_k^v(\lambda, u, t) \leq 0 \text{ for all } t \in [t_0, t_f], \quad k \in \{1, \dots, M\}, \\ y \in C([t_0, t_f]; Y), \quad u \in U_{[t_0, t_f]}, \end{array} \right\} \quad (24)$$

see, for example, [9, 6]. Further, we introduce the sets

$$F_v(\lambda) = \{\mu^* \in (C([t_0, t_f])_+^*)^M : h_v(\lambda, \mu^*) > -\infty\} \quad (25)$$

and

$$G(\lambda) = \left\{ \rho \in \prod_{v \in V_{[t_0, t_f]}} F_v(\lambda) : \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \rho_v) \in \mathbb{R} \right\}, \quad (26)$$

and, for $(r, \rho) \in \mathbb{R} \times G(\lambda)$, we define the projection $\pi(r, \rho) = r$. Finally, we introduce the following maximization problem as the dual problem of (7)

$$\left. \begin{array}{l} \text{maximize } \pi(r, \rho) \text{ subject to} \\ \rho \in G(\lambda), \quad r \in \mathbb{R}, \\ r \leq h_v(\lambda, \rho_v) \text{ for all } v \in V_{[t_0, t_f]}. \end{array} \right\} \quad (27)$$

The optimal value of this dual problem is

$$\Delta(\lambda) = \sup_{\rho \in G(\lambda)} \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \rho_v). \quad (28)$$

Now we state a strong duality result. For the convenience of the reader we also present a complete proof. Note however that Theorem 2 can also be deduced from Ky Fan's minimax theorem in [4].

Theorem 2 (Strong duality). *The constraint qualification (CQ) implies that*

$$\nu(\lambda) = \Delta(\lambda), \text{ for all } \lambda \in B(\bar{\lambda}), \quad (29)$$

where $\nu(\lambda)$ is the optimal value of (7) with fixed initial state.

Proof. Choose $\rho \in G(\lambda)$. Then convex weak duality implies that for all $v \in V_{[t_0, t_f]}$ we have

$$h_v(\lambda, \rho_v) \leq \nu^v(\lambda). \quad (30)$$

Thus

$$\inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \rho_v) \leq \inf_{v \in V_{[t_0, t_f]}} \nu^v(\lambda) = \nu(\lambda). \quad (31)$$

This implies that

$$\Delta(\lambda) = \sup_{\rho \in G(\lambda)} \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \rho_v) \leq \nu(\lambda), \quad (32)$$

that is, we have shown the weak duality. Further, due to (CQ) and convex strong duality from (23), for each $v \in V_{[t_0, t_f]}$, we can choose some $\mu_v^* \in (C([t_0, t_f])_+^*)^M$ such that

$$h_v(\lambda, \mu_v^*) = \nu^v(\lambda). \quad (33)$$

Define $\rho^* = (\mu_v^*)_{v \in V_{[t_0, t_f]}}$. Then $\rho^* \in G(\lambda)$. This yields

$$\begin{aligned} \Delta(\lambda) &= \sup_{\rho \in G(\lambda)} \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \rho_v) \\ &\geq \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda, \mu_v^*) = \inf_{v \in V_{[t_0, t_f]}} \nu^v(\lambda) = \nu(\lambda). \end{aligned} \quad (34)$$

Hence the strong duality follows. \square

Based upon the above duality concept, we can now show the Lipschitz continuity of the optimal value function in a neighborhood of $\bar{\lambda}$. To this end, we introduce for any $\varepsilon \geq 0$ the set of ε -optimal points

$$\begin{aligned} P(\lambda, \varepsilon) &= \{u \in U_{[t_0, t_f]} : \text{there exists } v \in V_{[t_0, t_f]} \text{ such that} \\ &\quad g_k^v(\lambda, u, t) \leq 0 \text{ for all } t \in [t_0, t_f], k = 1, \dots, M, \\ &\quad \varphi(\lambda, y(u, v), u, v) \leq \nu(\lambda) + \varepsilon\} \end{aligned} \quad (35)$$

and we set $H(\lambda, \varepsilon) = \{\rho \in G(\lambda) : \inf_{v \in V} h_v(\lambda, \rho_v) \geq \nu(\lambda) - \varepsilon\}$.

Lemma 2. *Under (CQ), the set*

$$\Omega(\bar{\lambda}) := \bigcup_{\lambda_1, \lambda_2 \in B(\bar{\lambda}), v \in V_{[t_0, t_f]}} \{\rho_v : \rho \in H(\lambda_1, |\lambda_1 - \lambda_2|^2)\} \quad (36)$$

is bounded.

Proof. Due to assumption (CQ), for all $v \in V_{[t_0, t_f]}$, we have the Slater point \bar{u}_v . Choose $\lambda_1, \lambda_2 \in B(\bar{\lambda})$ and $\rho \in H(\lambda_1, |\lambda_1 - \lambda_2|^2)$. Then $\inf_{v \in V_{[t_0, t_f]}} h_v(\lambda_1, \rho_v) \geq \nu(\lambda_1) - |\lambda_1 - \lambda_2|^2$. Thus by definition of h_v , for all $v \in V_{[t_0, t_f]}$, we have that $\mathcal{L}_v(\lambda_1, \bar{u}_v, \rho_v) \geq h_v(\lambda_1, \rho_v) \geq \nu(\lambda_1) - |\lambda_1 - \lambda_2|^2$. By definition of \mathcal{L}_v , this implies

$$\varphi(\lambda_1, y(\bar{u}_v, v), \bar{u}_v, v) + \sum_{k=1}^M \int_{t_0}^{t_f} g_k^v(\lambda_1, \bar{u}_v, s) d(\rho_v)_k(s) \geq \nu(\lambda_1) - |\lambda_1 - \lambda_2|^2. \quad (37)$$

Now using that

$$g_k^v(\lambda_1, \bar{u}_v, t) \leq -\omega < 0 \text{ for all } t \in [t_0, t_f], k \in \{1, \dots, M\}, \quad (38)$$

we can divide by $-\omega < 0$ and obtain due to (19)

$$\begin{aligned} \sum_{k=1}^M \int_{t_0}^{t_f} 1 d(\rho_v)_k(s) &\leq \frac{\nu(\lambda_1) - |\lambda_1 - \lambda_2|^2 - \varphi(\lambda_1, y(\bar{u}_v, v), \bar{u}_v, v)}{-\omega} \\ &= \frac{|\lambda_1 - \lambda_2|^2 + \varphi(\lambda_1, y(\bar{u}_v, v), \bar{u}_v, v) - \nu(\lambda_1)}{\omega} \\ &\leq \frac{|\lambda_1 - \lambda_2|^2 + \varphi(\lambda_1, y(\bar{u}_v, v), \bar{u}_v, v) - \underline{\alpha}}{\omega} \\ &\leq \frac{|\lambda_1 - \lambda_2|^2 + \sup_{v \in V_{[t_0, t_f]}} \sup_{\lambda \in B(\bar{\lambda})} \varphi(\lambda, y(\bar{u}_v, v), \bar{u}_v, v) - \underline{\alpha}}{\omega}. \end{aligned}$$

Due to (18) this yields the assertion. \square

Lemma 3. *Suppose that (CQ) holds. Then for all $\lambda_1, \lambda_2 \in B(\bar{\lambda})$ we have*

$$\nu(\lambda_1) - \nu(\lambda_2) \geq -\underline{C} |\lambda_1 - \lambda_2| \quad (39)$$

for some \underline{C} in \mathbb{R} .

Proof. Let $\lambda_1, \lambda_2 \in B(\bar{\lambda})$ be given. Choose a solution $u \in P(\lambda_1, |\lambda_1 - \lambda_2|^2)$ and $\tilde{v} \in V_{[t_0, t_f]}$ with $g_j^{\tilde{v}}(\lambda_1, u, t) \leq 0$ for all $t \in [t_0, t_f]$, $j = 1, \dots, M$, $\varphi(\lambda_1, y(u, \tilde{v}), u, \tilde{v}) \leq \nu(\lambda_1) + |\lambda_1 - \lambda_2|^2$ and $\bar{\rho} \in H(\lambda_2, |\lambda_1 - \lambda_2|^2)$. Then we have

$$\begin{aligned} \nu(\lambda_1) - \nu(\lambda_2) &\geq \varphi(\lambda_1, y(u, \tilde{v}), u, \tilde{v}) - \inf_{v \in V_{[t_0, t_f]}} h_v(\lambda_2, \bar{\rho}_v) - 2|\lambda_1 - \lambda_2|^2 \\ &\geq \varphi(\lambda_1, y(u, \tilde{v}), u, \tilde{v}) - h_{\tilde{v}}(\lambda_2, \bar{\rho}_{\tilde{v}}) - 2|\lambda_1 - \lambda_2|^2 \\ &\geq \varphi(\lambda_1, y(u, \tilde{v}), u, \tilde{v}) - \mathcal{L}_{\tilde{v}}(\lambda_2, u, \bar{\rho}_{\tilde{v}}) - 2|\lambda_1 - \lambda_2|^2 \\ &\geq \varphi(\lambda_1, y(u, \tilde{v}), u, \tilde{v}) + \sum_{j=1}^M \int_{t_0}^{t_f} g_j^{\tilde{v}}(\lambda_1, u, s) d\bar{\rho}_{\tilde{v}}(s) \\ &\quad - \mathcal{L}_{\tilde{v}}(\lambda_2, u, \bar{\rho}_{\tilde{v}}) - 2|\lambda_1 - \lambda_2|^2 \\ &= \mathcal{L}_{\tilde{v}}(\lambda_1, u, \bar{\rho}_{\tilde{v}}) - \mathcal{L}_{\tilde{v}}(\lambda_2, u, \bar{\rho}_{\tilde{v}}) - 2|\lambda_1 - \lambda_2|^2 \\ &\geq - \left[L_{\varphi}(|y(u, \tilde{v})|, |u|) \right. \\ &\quad \left. + M L_g(|u|) \int_{t_0}^{t_f} d\bar{\rho}_{\tilde{v}}(s) + 2|\lambda_1 - \lambda_2| \right] |\lambda_1 - \lambda_2|. \end{aligned}$$

Due to (CQ), the set $\bar{S}(y_0)$ from (20) is bounded. Thus our assumptions imply that the set $\{y(\hat{u}, \hat{v}) : (\hat{u}, \hat{v}) \in \bar{S}(y_0)\}$ is bounded (see (9)). Due to Lemma 2, the set $\Omega(\bar{\lambda})$ is also bounded. Since L_{φ} and L_g are continuous this allows us to define the real number

$$\begin{aligned} \tilde{C} &= \sup_{(\hat{u}, \hat{v}) \in \bar{S}(y_0)} L_{\varphi}(|y(\hat{u}, \hat{v})|, |\hat{u}|) \\ &\quad + M L_g(|\hat{u}|) \sup_{\hat{\rho}_w \in \Omega(\bar{\lambda})} \int_{t_0}^{t_f} d\hat{\rho}_w(s) + 2 \sup_{\lambda_1, \lambda_2 \in B(\bar{\lambda})} |\lambda_1 - \lambda_2|. \end{aligned} \quad (40)$$

Due to the definition of $P(\lambda_1, |\lambda_1 - \lambda_2|^2)$ we have $(u, \tilde{v}) \in \bar{S}(y_0)$. Moreover, we have $\bar{\rho}_{\tilde{v}} \in \Omega(\bar{\lambda})$. Hence we have

$$\nu(\lambda_1) - \nu(\lambda_2) \geq -\tilde{C} |\lambda_1 - \lambda_2|$$

and the assertion follows with $\underline{C} = \tilde{C}$. \square

Similarly as in Lemma 3, by interchanging the roles of λ_1 and λ_2 , and with the choice $\bar{C} = \tilde{C}$ with \tilde{C} as defined in (40) we can prove the following Lemma:

Lemma 4. *Suppose that (CQ) holds. Then, for all $\lambda_1, \lambda_2 \in B(\bar{\lambda})$, we have*

$$\nu(\lambda_1) - \nu(\lambda_2) \leq \bar{C} |\lambda_1 - \lambda_2|, \quad (41)$$

for some \bar{C} in \mathbb{R} .

The above analysis implies our main result about the Lipschitz continuity of the optimal value as a function of the parameter λ .

Theorem 3. *Under the Assumptions 1–3, for any $\bar{\lambda} \in \Lambda$ and a bounded neighborhood $B(\bar{\lambda}) \subset \Lambda$ satisfying the constraint qualification (CQ) it holds*

$$|\nu(\lambda_1) - \nu(\lambda_2)| \leq \tilde{C} |\lambda_1 - \lambda_2| \quad \text{for all } \lambda_1, \lambda_2 \in B(\bar{\lambda}) \quad (42)$$

with \tilde{C} as defined in (40), that is, the optimal value function ν is Lipschitz continuous in a neighborhood of $\bar{\lambda}$ with Lipschitz constant \tilde{C} .

Proof. The result follows from combining the proofs of Lemma 3 and 4. \square

4. JOINT PERTURBATIONS

In this section, we study the joint local Lipschitz continuity of the value function ν with respect to λ acting on the initial data, the constraints and the costs. We consider the mixed-integer optimal control problem (7). In contrast to Section 3 the initial state $y_0(\lambda)$ depends on λ . Also, the constraints and the objective function depend on λ . The result is obtained by combining Theorem 1 and 3.

Theorem 4. *Under the Assumptions 1–3, for any $\bar{\lambda} \in \Lambda$, a bounded neighborhood $B(\bar{\lambda}) \subset \Lambda$ let L_0, L_φ be constants such that (11) and (12) hold as in Theorem 1. Further, suppose that (CQ) holds in the sense that (17) is satisfied and $\cup_{y_0 \in Y_0} \bar{S}(y_0)$ is bounded with $\bar{S}(y_0)$ from (20) and $Y_0 = \{y_0(\lambda) : \lambda \in B(\bar{\lambda})\}$. Then, there exists a constant L_ν such that*

$$|\nu(\lambda_1) - \nu(\lambda_2)| \leq L_\nu |\lambda_1 - \lambda_2|, \quad \text{for all } \lambda_1, \lambda_2 \in B(\bar{\lambda}), \quad (43)$$

where $\nu(\lambda)$ is the optimal value of (7) as defined in (8).

Proof. In this proof, for $\lambda \in B(\bar{\lambda})$ and $y_0 \in Y$, we use the notation

$$\begin{aligned} \nu(\lambda, y_0) = \inf \{ & \varphi(\lambda, y, u, v) : \\ & \dot{y}(t) = A y(t) + f(t, y(t), u(t), v(t)), \quad t \in (t_0, t_f) \text{ a.e., } y(0) = y_0, \\ & g_k^v(\lambda, u, t) \leq 0 \text{ for all } t \in [t_0, t_f], \quad k = 1, \dots, M, \\ & y \in C([t_0, t_f]; Y), \quad u \in U_{[t_0, t_f]}, \quad v \in V_{[t_0, t_f]} \}. \end{aligned} \quad (44)$$

Due to (11) and (12) the set Y_0 is bounded by the constant K from Theorem 1 and for all $y_0 \in Y_0$, $v \in V_{[t_0, t_f]}$, $\lambda \in B(\bar{\lambda})$ we have the upper bound

$$\begin{aligned} \varphi(\lambda, y(y_0, \bar{u}_v, v), \bar{u}_v, v) &\leq \varphi(\bar{\lambda}, y(y_0(\bar{\lambda}), \bar{u}_v, v), \bar{u}_v, v) \\ &+ L_\varphi (|y(y_0, \bar{u}_v, v), \bar{u}_v, v) - y(y_0(\bar{\lambda}), \bar{u}_v, v), \bar{u}_v, v| + |\lambda - \bar{\lambda}|). \end{aligned}$$

Moreover, from Lemma 1, we obtain $|y(y_0, \bar{u}_v, v)| \leq C(t_f)(1 + K)$. This implies (18). Using similar arguments, we get a lower bound

$$\nu(\lambda, y_0) \geq \inf_{y_0 \in Y_0} \inf_{v \in V_{[t_0, t_f]}} \inf_{\lambda \in B(\bar{\lambda})} \varphi(\lambda, y(y_0, \bar{u}_v, v), \bar{u}_v, v) =: \underline{\alpha} > -\infty.$$

This implies (19) with $\underline{\alpha}$ independent of λ . Thus Assumptions 4 holds for all $y_0 \in Y_0$ and the proof of Lemma 2 shows that the bound of the set $\Omega(\bar{\lambda})$ is independent of y_0 . Due to (11) the function L_φ in Assumptions 3 is constant and \tilde{C} from (40) reduces to

$$\begin{aligned} \tilde{C} = L_\varphi + M \sup_{y_0 \in Y_0} \sup_{(\hat{u}, \hat{v}) \in \bar{S}(y_0)} L_g(|\hat{u}|) \sup_{\hat{\rho}_w \in \Omega(\bar{\lambda})} \int_{t_0}^{t_f} d\hat{\rho}_w(s) \\ + 2 \sup_{\lambda_1, \lambda_2 \in B(\bar{\lambda})} |\lambda_1 - \lambda_2| < \infty. \end{aligned}$$

Now, let $\lambda_1, \lambda_2 \in B(\bar{\lambda})$. From Theorem 1 with λ_1 as first argument of ν fixed we get $|\nu(\lambda_1, y_0(\lambda_1)) - \nu(\lambda_1, y_0(\lambda_2))| \leq \hat{L}_\nu |\lambda_1 - \lambda_2|$ with \hat{L}_ν given by (14). From Theorem 3 with λ_2 as an argument of y_0 fixed we get $|\nu(\lambda_1, y_0(\lambda_2)) - \nu(\lambda_2, y_0(\lambda_2))| \leq \tilde{C} |\lambda_1 - \lambda_2|$. Thus we obtain the inequality

$$\begin{aligned} |\nu(\lambda_1, y_0(\lambda_1)) - \nu(\lambda_2, y_0(\lambda_2))| \\ \leq |\nu(\lambda_1, y_0(\lambda_1)) - \nu(\lambda_1, y_0(\lambda_2))| + |\nu(\lambda_1, y_0(\lambda_2)) - \nu(\lambda_2, y_0(\lambda_2))| \\ \leq (\hat{L}_\nu + \tilde{C}) |\lambda_1 - \lambda_2| \end{aligned}$$

and (43) follows with $L_\nu = \hat{L}_\nu + \tilde{C}$. \square

5. EXAMPLE

We discuss an academic application concerning the optimal positioning of an actuator motivated from applications in thermal manufacturing [12, 11].

Example 2. Suppose that $\Omega \subset \mathbb{R}^2$ is a bounded domain containing two non-overlapping control domains ω_1 and ω_2 . For simplicity, we assume that the boundaries of all these domains $\partial\Omega$, $\partial\omega_1$ and $\partial\omega_2$ are smooth. Let $\varepsilon, \delta > 0$ be two given parameters, let χ_{ω_i} denote the characteristic function of ω_i , let $v|_{[t_1, t_2]^+}$ denote the restriction of v to the non-negative part of the interval $[t_1, t_2]$ and let Δy denote the Laplace operator. For a time horizon with $t_f - t_0 > \delta$, we consider the optimal control problem

$$\begin{aligned} & \text{minimize} \quad \int_{t_0}^{t_f} \int_{\Omega} |y(t, x) - \hat{y}(t, x)|^2 dx dt + \int_{t_0}^{t_f} |u(t)|^2 dt \\ & y_t - \Delta y + v(t)u(t)\chi_{\omega_1} + (1 - v(t))u(t)\chi_{\omega_2} = 0 \quad \text{on } (t_0, t_f) \times \Omega \\ & y = 0 \quad \text{on } (t_0, t_f) \times \partial\Omega \\ & y = \bar{y} \quad \text{on } \{t_0\} \times \Omega \\ & v(t) \in \mathcal{V} = \{0, 1\} \quad \text{on } [t_0, t_f] \\ & u(t) \in \begin{cases} [0, 1 + \varepsilon] & \text{if } v|_{[t-\delta, t]^+} \equiv 1 \text{ or } v|_{[t-\delta, t]^+} \equiv 0 \text{ a. e. on } [t - \delta, t]^+ \\ [0, \varepsilon] & \text{else.} \end{cases} \end{aligned}$$

The combination of the actuator and constraints in this problem model that the continuous control u is restricted to a small uncontrollable disturbance ε for a dwell-time period of length δ whenever a decision was taken to change the control region ω_1 to ω_2 or vice versa while the goal is to steer the initial state $\bar{y} \in L^2(\Omega)$ as close as possible to a desired state $\hat{y} \in C([t_0, t_f]; L^2(\Omega))$. We consider a perturbation $\lambda = (\bar{y}, \varepsilon, \hat{y})$, i. e., a joint perturbation of initial data, the disturbance and the tracking target.

We can consider this problem in the abstract setting with $Y = L^2(\Omega)$, $U = \mathbb{R}$, $U_{[t_0, t_f]} = L^\infty(t_0, t_f)$

$$\begin{aligned} Ay &= \Delta y, \quad y \in D(A) = H^2(\Omega) \cap H_0^1(\Omega), \\ f(y, u, v) &= f(u, v) = -(vu\chi_{\omega_1} + (1 - v)u\chi_{\omega_2}), \\ \varphi(\lambda, y, u, v) &= \int_{t_0}^{t_f} \int_{\Omega} |y(t, x) - \hat{y}(t, x)|^2 dx dt + \int_{t_0}^{t_f} |u(t)|^2 dt, \end{aligned}$$

defining

$$\bar{u}_\varepsilon^v(t) = \begin{cases} 1 + \varepsilon & \text{if } v|_{[t-\delta, t]^+} \equiv 1 \text{ or } v|_{[t-\delta, t]^+} \equiv 0 \text{ a. e. on } [t - \delta, t]^+ \\ \varepsilon & \text{else,} \end{cases}$$

and setting $M = 2$ and, for all $v \in V_{[t_0, t_f]}$

$$\begin{aligned} g_1^v(\lambda, u, t) &= \text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_\varepsilon^v(s)), \\ g_2^v(\lambda, u, t) &= \text{ess sup}_{s \in [t_0, t_f]} (-u(s)). \end{aligned}$$

Here the $g_i^v(\lambda, u, \cdot)$ are constant with respect to t and hence continuous as functions of t . Moreover, the maps $u \mapsto g_i^v(\lambda, u, \cdot)$ are continuous in $L^\infty(t_0, t_f)$. The objective function is convex with respect to (y, u) and also the maps $u \mapsto g_i^v(\lambda, u, t)$ are convex. Let $\varepsilon_1, \varepsilon_2 > 0$ be such that without restriction we have

$\text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_{\varepsilon_1}^v(s)) \geq \text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_{\varepsilon_2}^v(s))$. Then we have

$$\begin{aligned}
& |g_1^v(\lambda_1, u, t) - g_1^v(\lambda_2, u, t)| \\
&= \text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_{\varepsilon_1}^v(s)) - \text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_{\varepsilon_2}^v(s)) \\
&= \text{ess sup}_{s \in [t_0, t_f]} (u(s) + \bar{u}_{\varepsilon_2}^v(s) - \bar{u}_{\varepsilon_1}^v(s) - \bar{u}_{\varepsilon_2}^v(s)) + \\
&\quad - \text{ess sup}_{s \in [t_0, t_f]} (u(s) - \bar{u}_{\varepsilon_2}^v(s)) \\
&\leq \text{ess sup}_{s \in [t_0, t_f]} |\bar{u}_{\varepsilon_2}^v(s) - \bar{u}_{\varepsilon_1}^v(s)| \\
&\leq |\varepsilon_2 - \varepsilon_1|.
\end{aligned}$$

It is well-known that $(A, D(A))$ is the generator of a strongly continuous (analytic) semigroup of contractions $\{T(t)\}_{t \geq 0}$ on Y , see, e.g., [14]. Also, Assumptions 1–3 are easily verified and it is easy to see that (11) and (12) hold. The constraint qualification (CQ) is satisfied with $\omega = \frac{\varepsilon}{2}$ and $\underline{\alpha} = 0$. The control constraints imply that the set \bar{S} is bounded in $U_{[t_0, t_f]} \times V_{[t_0, t_f]}$ independently of the initial state y_0 . Hence we can conclude from Theorem 4 that the optimal value function ν is locally Lipschitz continuous jointly as a function of $\lambda = (\bar{y}, \varepsilon, \hat{y})$.

6. CONCLUSION

We have studied the optimal value function for control problems on Banach spaces that involve both continuous and discrete control decisions. For control systems of a semilinear type subject to control constraints, we have shown that the optimal value depends locally Lipschitz continuously on perturbations of the initial data and costs under natural assumptions. For problems consisting of linear systems on a Banach space subject to convex control inequality constraints, we have shown that the optimal value of convex cost functions depend locally Lipschitz continuously on Lipschitz continuous perturbations of the costs and the constraints under a Slater-type constraint qualification. The result has been obtained by proving a strong duality for an appropriate dual problem.

By a combination of the above results we have for the linear, convex case obtained local Lipschitz continuity jointly for parametric initial data, control constraints and cost functions. The Example 1 shows that this result is sharp in the sense that we can, in general, not expect much more regularity than we have proved.

Our analysis currently does not address the stability of the optimal control under perturbations. This is an interesting direction for future work.

ACKNOWLEDGEMENTS

This work was supported by the DFG grant CRC/Transregio 154, projects C03 and A03. The authors thank the reviewers and the editors from *Math. Control Signals Syst.* for the constructive suggestions. The final publication is available at link.springer.com and [doi:10.1007/s00498-016-0183-4](https://doi.org/10.1007/s00498-016-0183-4).

REFERENCES

- [1] Viorel Barbu and Giuseppe Da Prato. *Hamilton-Jacobi equations in Hilbert spaces*, volume 86 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1983.
- [2] Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, and Sanjoy K. Mitter. *Representation and control of infinite-dimensional systems. Vol. 1*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1992.
- [3] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000.
- [4] J. M. Borwein and D. Zhuang. On fan's minimax theorem. *Math Program*, 26:232–234, 1986.
- [5] Piermarco Cannarsa and Halina Frankowska. Value function and optimality conditions for semilinear control problems. *Appl Math Optim*, 26(2):139–169, 1992.

- [6] Ivar Ekeland and Thomas Turnbull. *Infinite-dimensional optimization and convexity*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1983.
- [7] Jacques Gauvin and François Dubeau. Differential properties of the marginal function in mathematical programming. *Math Programming Stud*, (19):101–119, 1982.
- [8] Matthias Gerdts. A variable time transformation method for mixed-integer optimal control problems. *Optimal Control Appl. Methods*, 27(3):169–182, 2006.
- [9] M. Gugat. One-sided derivatives for the value function in convex parametric programming. *Optimization*, 28(3-4):301–314, 1994.
- [10] M. Gugat. Parametric disjunctive programming: one-sided differentiability of the value function. *J Optim Theory Appl*, 92(2):285–310, 1997.
- [11] Falk M. Hante and Sebastian Sager. Relaxation methods for mixed-integer optimal control of partial differential equations. *Comput Optim Appl*, 55(1):197–225, 2013.
- [12] Orest V. Iftime and Michael A. Demetriou. Optimal control of switched distributed parameter systems with spatially scheduled actuators. *Automatica J. IFAC*, 45(2):312–323, 2009.
- [13] B. S. Mordukhovich, N. M. Nam, and N. D. Yen. Subgradients of marginal functions in parametric mathematical programming. *Math Program*, 116(1-2, Ser. B):369–396, 2009.
- [14] A. Pazy. *Semigroups Of Linear Operators And Applications To Partial Differential Equations*. Applied Mathematical Sciences Series, Springer-Verlag, New York, 1983.
- [15] Fabian Rüffler and Falk M. Hante. Optimal switching for hybrid semilinear evolutions. *Non-linear Anal Hybrid Syst*, 22:215–227, 2016.
- [16] Sebastian Sager. Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control. *J Process Control*, 19(8):1238–1247, 2009.
- [17] Ramanarayan Vasudevan, Humberto Gonzalez, Ruzena Bajcsy, and S. Shankar Sastry. Consistent approximations for the optimal control of constrained switched systems—part 1: A conceptual algorithm. *SIAM J Control Optim*, 51(6):4463–4483, 2013.
- [18] Ramanarayan Vasudevan, Humberto Gonzalez, Ruzena Bajcsy, and S. Shankar Sastry. Consistent approximations for the optimal control of constrained switched systems—part 2: An implementable algorithm. *SIAM J Control Optim*, 51(6):4484–4503, 2013.
- [19] A. C. Williams. Marginal values in mixed integer linear programming. *Math Program*, 44(1, (Ser. A)):67–75, 1989.
- [20] Feng Zhu and Panos J. Antsaklis. Optimal control of hybrid switched systems: A brief survey. *Discrete Event Dyn S*, 25(3):345–364, 2015.
- [21] Enrique Zuazua. Switching control. *J Eur Math Soc*, 13:85–117, 2011.